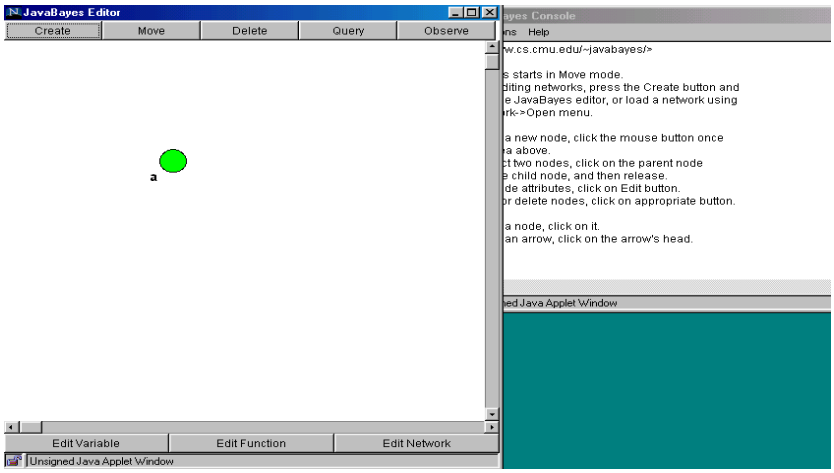
Java Bayes Laboratory

To access the Java Bayes,
Type in the following on your browser:
http://www-2.cs.cmu.edu/~javabayes/Home/

Once you get to the Java Bayes Page, click on the applet and the JavaBayes editor and the JavaBayes console should appear on your screen.

To create new nodes, select the:

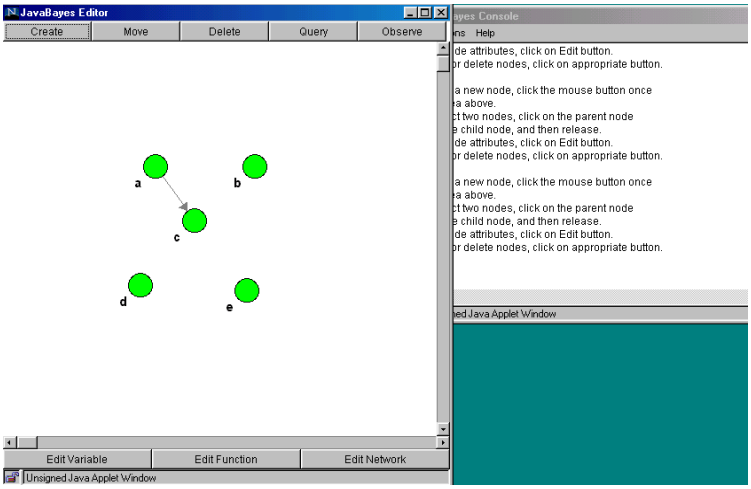
1. 'Create' from the top line of the JavaBayes editor
2. Move the cursor to a particular position on the screen, and click mouse key.
3. Continue to do this until you have all the nodes you need



To connect the nodes to each other:

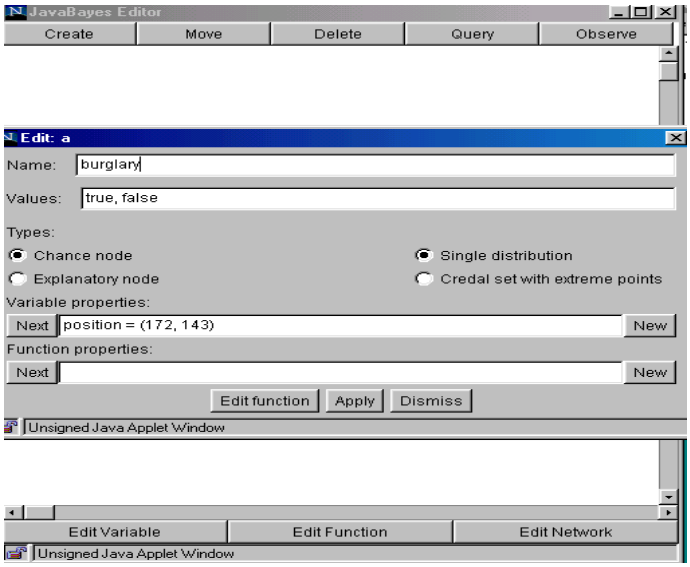
1. While still in the create mode, place cursor on one of the nodes and click mouse button
2. Drag mouse to the connecting node and release mouse button. Make sure that you get the curser in the circle.

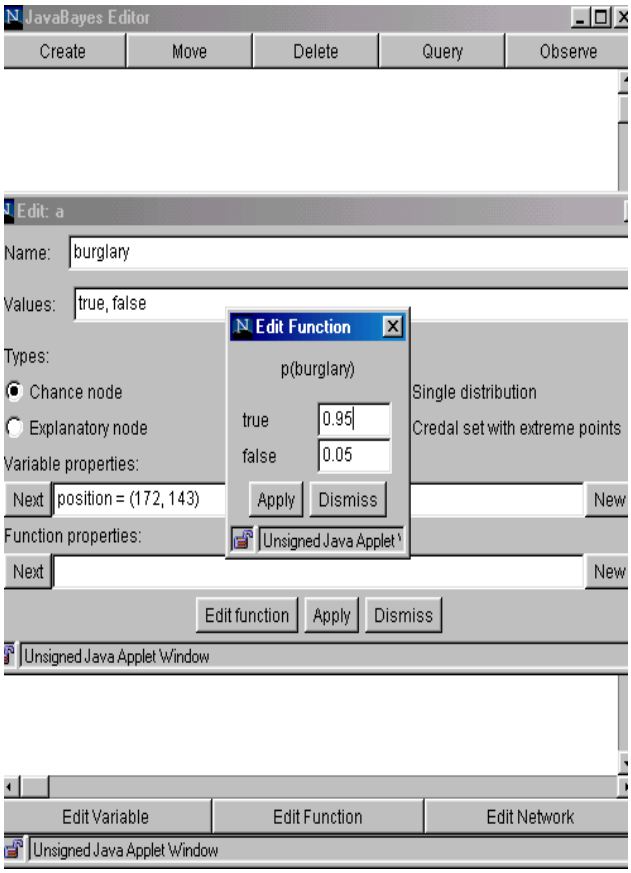
The following should appear on your screen.



To add values to the nodes:

1. Click on Edit Variable selection at the bottom of the JavaBayes Editor
2. Click on Any One of the nodes:
   a. To add a name:
   Type in the appropriate name in the name field and click “apply”
   b. To add variable values:
   Select “Edit Function” at the bottom of the editor and then enter the values.





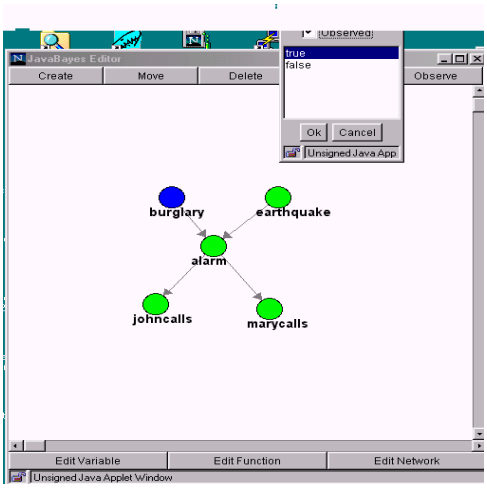
For the nodes with multiple inputs, you will have to enter multiple values, one for each of the different joint probabilities…..

See below: Make sure that you recognize that they are going horizontal rather than vertical (as in the book).

Querying the system is a two step process,

1. You need to identify which variables have been "observed" (to be either true or false). For example, if we are interested in the P (John Calls | Burglary), and you need to tell the system that what those values are (example: Burglary has been observed to be true.)
2. To do this, select "observe" from the JavaBayes menu
3. Click on the "burglary" node.
4. Select true and indicate that it has been observed.

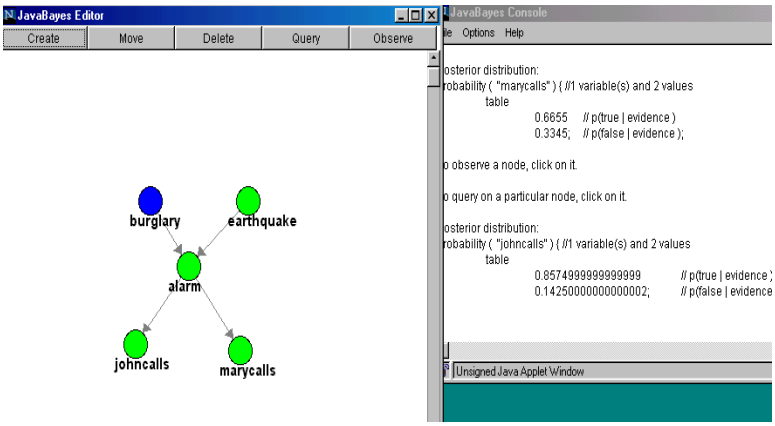
See below:



Now you can query the system by:

1. Selecting Query from the JavaBayes Editor menu
2. Clicking on the node of interest (that is, john calls).

The following should appear in the Console Window.



SCENARIO:

Suppose you are working for a financial institution and you are asked to implement a fraud detection system. You plan to use the following information: When the card holder is travelling abroad, fraudulent transactions are more likely since tourists are prime targets for thieves. More precisely, 1% of transactions are fraudulent when the card holder is travelling; whereas only 0.2% of the transactions are fraudulent when she is not travelling. On average, 5% of all transactions happen while the card holder is travelling. If a transaction is fraudulent, then the likelihood of a foreign purchase increases, unless the card holder happens to be travelling. More precisely, when the card holder is not travelling, 10% of the fraudulent transactions are foreign purchases where as only 1% of the legitimate transactions are foreign purchases. On the other hand, when the card

holder is travelling, then 90% of the transactions are foreign purchases regardless of the legitimacy of the transactions.

Purchases made over the internet are more likely to be fraudulent. This is especially true for card holders who don't own any computer. Currently, 60% of the population owns a computer and for those card holders, 1% of their legitimate transactions are done over the internet, however this percentage increases to 2% for fraudulent transactions. For those who don't own any computer, a mere 0.1% of their legitimate transactions is done over the internet, but that number increases to 1.1% for fraudulent transactions. Unfortunately, the credit card company doesn't know whether a card holder owns a computer, however it can usually guess by verifying whether any of the recent transactions involve the purchase of computer related accessories. In any given week, 10% of those who own a computer purchase with their credit card at least one computer related item as opposed to just 0.1% of those who don't own any computer.

A. Construct a Bayes Network that your fraud detection system can use to indentify fraudulent transactions. Show the graph defining the network and the Conditional Probability Tables associated with each node in the graph. This network should encode the information stated above. Your network should contain exactly six nodes, corresponding to the following binary random variables:

_ OC – card holder owns a computer.
_ Fraud – current transaction is fraudulent.
_ Trav – card holder is currently travelling.
_ FP – current transaction is a foreign purchase.
_ IP – current purchase is an internet purchase.
_ CRP – a computer related purchase was made in the past week.

The arcs defining your Bayes Network should accurately capture the probabilistic dependencies between these variables.

B. What is the prior probability (i.e., before we search for previous computer related purchases and before we verify whether it is a foreign and/or an internet purchase) that the current transaction is a fraud? What is the probability that the current transaction is a fraud once we have verified that it is a foreign transaction, but not an internet purchase and that the card holder purchased computer related accessories in the past week?

C. After computing those probabilities, the fraud detection system raises a flag and recommends that the card holder be called to confirm the transaction. An agent calls at the domicile of the card holder but she is not home. Her spouse confirms that she is currently out of town on a business trip. How does the probability of a fraud changes based on this new piece of information?

D. Suppose you are not a very honest employee and you just stole a credit card. You know that the fraud detection system uses the Bayes net designed earlier but you still want to make an important purchase over the internet. What can you do prior to your internet purchase to reduce the risk that the transaction will be rejected as a possible fraud?